GRADIENT AND OPTIMIZATION

The gradient is closely related to optimization.

Prop: If $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ is differentiable at $p \in D$, then $D_{\vec{u}} f(p)$ attains its maximum in the direction $\nabla f(p)$.

pf: By the gradient formula for directional derivatives:

$$D_{\vec{u}} f(p) = \vec{\nabla} f(p) \cdot \vec{u} = |\vec{\nabla} f(p)||\vec{u}| \cos(\theta) = |\vec{\nabla} f(p)| \cos(\theta)$$

which is maximized when $\theta = 0$. ///

Ex: In which direction is $D_{\vec{u}} f(p)$ maximized for...

① $f(x,y) = 4y\sqrt{x}$ $\quad p = (4,1)$

Sol: $\nabla f = 4\langle \frac{1}{2}x^{-1/2}y, \sqrt{x} \rangle$.

$\therefore$ we maximize $D_{\vec{u}} f(p)$ for $\vec{u} = \nabla f(p) = 4\langle \frac{1}{4}, 2 \rangle = \langle 1, 8 \rangle$ ///

② $f(x,y,z) = \dfrac{x}{y+z}$ $\quad p = (8,1,3)$

Sol: $\nabla f = \left\langle \frac{1}{y+z}, -\frac{x}{(y+z)^2}, -\frac{x}{(y+z)^2} \right\rangle$, so maximize $D_{\vec{u}} f(p)$ at

$$\nabla f(p) = \left\langle \frac{1}{1+3}, -\frac{8}{(1+3)^2}, -\frac{8}{(1+3)^2} \right\rangle = \left\langle \frac{1}{4}, -\frac{8}{16}, -\frac{8}{16} \right\rangle = \frac{1}{4}\langle 1, -2, -2 \rangle$$ ///

③ $f(x,y,z) = x\arctan(z) - y$ $\quad p = (0,0,0)$

Sol: $\nabla f = \left\langle \arctan(z), -1, \frac{x}{1+z^2} \right\rangle$, so maximize $D_{\vec{u}} f(p)$

in direction of $\nabla f(0,0,0) = \langle 0, -1, 0 \rangle$. ///

Optimization

Goal: Extend local optimization techniques to functions of several variables. (NB: Assume f is diff)

Terminology: Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ and $p \in D$.

1. Function f has a local max at p when for all x near p we have $f(p) \geq f(x)$.
2. Function f has an absolute max at p when for all $x \in D$ we have $f(p) \geq f(x)$.
3. Local and absolute minima are defined similarly (i.e. just flip the inequalities ...)
4. A critical point of f is a point $p \in D$ where either

☆ $$\vec{\nabla} f(p) = \vec{0} \quad \text{or} \quad \vec{\nabla} f(p) \text{ is undefined.}$$

Ex: Compute all critical points of $f(x,y) = xy - 2x - 2y - x^2 - y^2$.

Sol: $\vec{\nabla} f(p)$ is always defined (f is polynomial).

$\vec{\nabla} f = \langle y - 2 - 2x, \ x - 2 - 2y \rangle$. So

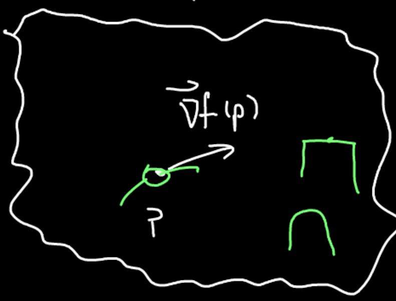
$$\vec{\nabla} f = 0 \iff \begin{cases} y - 2 - 2x = 0 \\ x - 2 - 2y = 0 \end{cases} \iff \begin{cases} y = 2 + 2x \\ x - 2 - 2(2+2x) = 0 \end{cases}$$

$$\iff \begin{cases} -3x - 6 = 0 \\ y = 2 + 2x \end{cases} \iff \begin{cases} x = -2 \\ y = -2 \end{cases}$$

$\therefore$ f has unique critical point $(-2, -2)$. $\blacksquare$

Prop (Fermat's Extremum Theorem): Local extrema are critical points.

Idea: If $\vec{\nabla} f(p) \neq \vec{0}$ at local maximum p, we could move a small amount in the direction of $\vec{\nabla} f(p)$ to attain an increase in the value of f in that direction (because $\vec{\nabla} f(p)$ maximizes $D_{\vec{u}} f(p)$, the instantaneous rate of change in direction $\vec{u}$).



pf: Suppose p is a local maximum point for f, and assume to the contrary that $\nabla f(p) \neq \vec{0}$. Define function $g: (-\varepsilon, \varepsilon) \to \mathbb{R}$ by formula $g(t) := f(\vec{p} + t\vec{u})$ where $\vec{u} = \frac{\nabla f(p)}{|\nabla f(p)|}$. By earlier work, we have

$$D_{\vec{u}} f(p) = \vec{\nabla} f(p) \cdot \frac{\vec{\nabla} f(p)}{|\vec{\nabla} f(p)|} = |\nabla f(p)| > 0.$$

Notice

$$D_{\vec{u}} f(p) = \lim_{h \to 0} \frac{f(\vec{p} + h\vec{u}) - f(\vec{p})}{h} = \lim_{h \to 0} \frac{g(0+h) - g(0)}{h} = g'(0).$$

As p is a local maximum of f, we have 0 is a local maximum of g (otherwise we could always find a larger value for g near 0, and g-values are f-values!). By the single-variable Extremum Theorem, g has a critical point at 0. Hence $0 = g'(0) = D_{\vec{u}} f(p) = |\nabla f(p)| > 0$.

But this is absurd! Hence $\nabla f(p) = \vec{0}$. ∎

Point: Optimize locally by computing critical points!

Question: Which critical points yield local extrema?

I.E. Is there a first/second derivative test analogue?

Answer: Yes... But it's complicated...

too complicated, don't use!

Prop (First Derivative Test): If p is a critical point and for all unit vectors $\vec{u}$ there is an $\varepsilon > 0$ such that for all $t \in (0, \varepsilon)$ we have...

1. $D_{\vec{u}}(\vec{p} + t\vec{u}) \geq 0$, then p is a local minimum.
2. $D_{\vec{u}}(\vec{p} + t\vec{u}) \leq 0$, then p is a local maximum.

Yikes... Let's not do that...

In fact, let's make two simplifying assumptions:

1. f has continous second partial derivatives.
2. $f = f(x,y)$ is a function of two variables

We define
$$D(a,b) := \begin{vmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{vmatrix} = (f_{xx} \cdot f_{yy} - f_{xy}^2)(a,b).$$

Prop (Second Derivative Test): If f has cts 2^{nd} partial derivatives at critical point $p = (a,b)$ and...

1. $D(a,b) > 0$ and $f_{xx}(a,b) > 0$, then (a,b) yields a loc. min.
2. $D(a,b) > 0$ and $f_{xx}(a,b) < 0$, then (a,b) yields a loc. max.
3. $D(a,b) < 0$, then (a,b) yields a saddle point. (one direction: $\cap$, another: $\cup$)

NB: • The test fails if $D(a,b) = 0$. (:|)

• A saddle point is a place where $f(x,y)$ locally looks like a hyperbolic paraboloid.

• Local extrema look locally like elliptic paraboloids.

Ex: Classify the critical points of the function

$$f(x,y) = xy - 2x - 2y - x^2 - y^2.$$

Sol: From earlier: Critical point: $(-2,-2)$.

Using the second derivative test:

$$D(-2,-2) = \begin{vmatrix} f_{xx}(-2,-2) & f_{xy}(-2,-2) \\ f_{yx}(-2,-2) & f_{yy}(-2,-2) \end{vmatrix} = ?$$

$$\nabla f(x,y) = \langle y - 2 - 2x,\ x - 2 - 2y \rangle$$

$$f_{xx} = -2, \quad f_{xy} = f_{yx} = 1, \quad f_{yy} = -2.$$

Hence $D(x,y) = \begin{vmatrix} -2 & 1 \\ 1 & -2 \end{vmatrix} = (-2)(-2) - 1 \cdot 1 = 3 > 0.$

$f_{xx}(x,y) = -2 < 0$. Hence $D(-2,-2) = 3 > 0$ and $f_{xx}(-2,-2) = -2 < 0$ yields that f has a local maximum at the critical point $(-2,-2)$. ☐

Ex: Classify critical points of $f(x,y) = 4 + x^3 + y^3 - 3xy$.

Sol: (Gradient): $\nabla f = \langle 3x^2 - 3y,\ 3y^2 - 3x \rangle$

(Crit Pts): $\nabla f(x,y) = \vec{0} \iff \begin{cases} 3x^2 - 3y = 0 \\ 3y^2 - 3x = 0 \end{cases}$

$\iff \begin{cases} y = x^2 \\ x = y^2 \end{cases} \iff \begin{cases} x = y^2 \\ y = x^2 = (y^2)^2 = y^4 \end{cases}$

$\iff \begin{cases} x = y^2 \\ y(y^3 - 1) = 0 \end{cases} \iff \begin{cases} x = y^2 \\ y = 0 \text{ or } y = 1 \end{cases}$

$\iff (x,y) \in \{ (0,0), (1,1) \}$

(Second order partials): $\dfrac{\partial^2 f}{(\partial x)^2} = 6x$, $\dfrac{\partial^2 f}{\partial x \partial y} = -3$, $\dfrac{\partial^2 f}{(\partial y)^2} = 6y$

(Compute $D(x,y)$): $D(x,y) = \begin{vmatrix} \frac{\partial^2 f}{(\partial x)^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{(\partial y)^2} \end{vmatrix} = f_{xx} \cdot f_{yy} - f_{xy}^2$

$= 6x \cdot 6y - (-3)^2 = 36xy - 9$

$= 9(4xy - 1)$

(Analyze CP's):

(0,0): $D(0,0) = 9(0 - 1) = -9 < 0$

Hence $(0,0)$ is a saddle point

(1,1): $D(1,1) = 9(4 \cdot 1 \cdot 1 - 1) = 9 \cdot 3 > 0$

$f_{xx}(1,1) = 6 \cdot 1 = 6 > 0$

Hence $(1,1)$ is a local minimum point. $\square$

Ex: Classify critical points of $f(x,y) = (x-y)(1-xy)$.

Sol: (Gradient): $\nabla f = (\nabla(x-y))(1-xy) + (x-y)(\nabla(1-xy))$

$$= (1-xy)\langle 1,-1\rangle + (x-y)\langle -y,-x\rangle$$

$$= \langle 1-xy-xy+y^2,\ xy-1-x^2+xy\rangle$$

$$= \langle 1-2xy+y^2,\ -(x^2-2xy+1)\rangle$$

(CPs): $\begin{cases} y^2-2xy+1=0 \\ x^2-2xy+1=0 \end{cases} \iff \begin{cases} x^2-2xy+1=0 \\ y^2-2xy+1=0 \\ x^2-y^2=0 \end{cases} \leftarrow$

$\iff \begin{cases} x=\pm y \\ x^2-2x^2+1=0 \text{ OR } x^2+2x^2+1=0 \\ x^2-2x^2+1=0 \text{ OR } x^2+2x^2+1=0 \end{cases} \iff \begin{cases} y=x \text{ or } y=-x \\ 1-x^2=0 \end{cases}$ ~~OR $1+3x^2=0$~~ $\iff \begin{cases} x=1 \text{ OR } x=-1 \\ y=x \text{ OR } y=-x \end{cases}$

$$\iff (x,y) \in \{(1,1),\ (1,-1),\ (-1,-1),\ (-1,1)\}$$

(Partial Der.vs): $f_{xx}(x,y) = -2y$, $f_{xy}(x,y) = -2x+2y$, $f_{yy} = +2x$

(compute $D(x,y)$): $D(x,y) = f_{xx}\cdot f_{yy} - f_{xy}^2 = (-2y)(2x) - (-2x+2y)^2$

$$= -4xy - 4(x+y)^2 = -4(x^2+3xy+y^2)$$

Analyze CP's:

$(1,1)$: $D(1,1) = -4(1^2+3+1^2) < 0$ $\therefore (1,1)$ is a saddle pt.

$(1,-1)$: $D(1,-1) = -4(1-3+1) = 4 > 0$, $f_{xx}(1,-1) = -(-1) > 0$

$\therefore (1,-1)$ is a local minimum point.

$(-1,-1)$: $D(-1,-1) = -4(1+3+1) < 0$ $\therefore (-1,-1)$ is a saddle pt.

$(-1,1)$: $D(-1,1) = -4(1-3+1) > 0$, $f_{xx}(-1,1) = -1 < 0$

$\therefore (-1,1)$ is a local maximum point. $\square$

Ex: Classify critical points of $f(x,y) = y(e^x - 1)$.

Sol: Gradient: $\nabla f = \langle e^x y, e^x - 1 \rangle$

Critical Points: $\nabla f = 0 \iff \begin{cases} e^x y = 0 \\ e^x - 1 = 0 \end{cases} \iff \begin{cases} y = 0 \\ e^x = 1 \end{cases} \iff \begin{cases} x = 0 \\ y = 0 \end{cases}$

$\therefore$ Crit pts $(0,0)$

Partial derivs: $f_{xx} = ye^x$, $f_{xy} = e^x$, $f_{yy} = 0$

Compute $D(0,0)$: $D(0,0) = 0 \cdot 0 - 1^2 = -1 < 0$

Analyze Crit. pts: $D(0,0) < 0$ yields $(0,0)$ is a saddle point. $\boxtimes$

Ex: Classify Critical points of $f(x,y) = xy + e^{-xy}$.

Sol: Gradient: $\nabla f = \langle y - ye^{-xy}, x - xe^{-xy} \rangle = (1 - e^{-xy}) \langle y, x \rangle$

Crit Pts: $\nabla f = 0 \iff 1 - e^{-xy} = 0$ OR $\begin{cases} y = 0 \\ x = 0 \end{cases}$

$\iff e^{-xy} = 0 \iff x = 0$ OR $y = 0$

Crit pts: $\{(x,0), (0,y) : x,y \in \mathbb{R}\}$

Second Order Partials: $f_{xx} = y^2 e^{-xy}$, $f_{xy} = 1 + (y^2 - 1) e^{-xy}$, $f_{yy} = x^2 e^{-xy}$

Compute $D(x,y)$: $D(x,y) = f_{xx} \cdot f_{yy} - f_{xy}^2 = x^2 y^2 e^{-2xy} - (1 + (y^2 - 1) e^{-xy})$

Analyze CPs: Cases: $x = 0$: $D(0,y) = -y^2 \begin{cases} < 0 & \text{if } y \neq 0 \\ = 0 & \text{if } y = 0 \end{cases}$

$(0,y)$ is a saddle point when $y \neq 0$.

$y = 0$: $D(x,0) = 0 - (1 + (0-1) \cdot 1) = 0$

Hence if $y = 0$, second deriv test yields no info. $\boxtimes$

Ex: Classify the critical points of $f(x,y) = xy\,e^{-\frac{x^2+y^2}{2}}$.

Sol:

Step 1 (Gradient):

$$\begin{aligned}
\nabla f(x,y) &= \nabla(xy)\cdot e^{-\frac{x^2+y^2}{2}} + xy\,\nabla\left(e^{-\frac{x^2+y^2}{2}}\right) \\
&= e^{-\frac{x^2+y^2}{2}}\,\nabla(xy) + xy\,e^{-\frac{x^2+y^2}{2}}\,\nabla\left(-\frac{x^2+y^2}{2}\right) \\
&= e^{-\frac{x^2+y^2}{2}}\left(\langle y, x\rangle - xy\,\langle x, y\rangle\right) \\
&= e^{-\frac{x^2+y^2}{2}}\,\langle y - x^2y,\ x - xy^2\rangle \\
&= e^{-\frac{x^2+y^2}{2}}\,\langle y(1-x)(1+x),\ x(1-y)(1+y)\rangle
\end{aligned}$$

Step 2 (critical points):

$$\nabla f(x,y) = 0 \iff \begin{cases} y(1-x)(1+x) = 0 \\ x(1-y)(1+y) = 0 \end{cases}$$

$$\iff \begin{cases} y = 0 \ \text{ OR } \ x = 1 \ \text{ OR } \ x = -1 \\ x = 0 \ \text{ OR } \ y = 1 \ \text{ OR } \ y = -1 \end{cases}$$

$$\iff (x,y) \in \{(0,0), (1,1), (1,-1), (-1,1), (-1,-1)\}$$

Hence we have critical points $(0,0)$, $(1,1)$, $(1,-1)$, $(-1,1)$, $(-1,-1)$.

Step 3 (Compute $D(x,y)$):

$$D(x,y) = \begin{vmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{yx}(x,y) & f_{yy}(x,y) \end{vmatrix}$$

$$\begin{aligned}
\frac{\partial^2 f}{(\partial x)^2} &= \frac{\partial}{\partial x}\left[y(1-x^2)e^{-\frac{x^2+y^2}{2}}\right] \\
&= y(1-x^2)\cdot -x\,e^{-\frac{x^2+y^2}{2}} + y\cdot -2x\,e^{-\frac{x^2+y^2}{2}} \\
&= -xy(3-x^2)\,e^{-\frac{x^2+y^2}{2}}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 f}{\partial x \partial y} &= \frac{\partial}{\partial y}\left[y(1-x^2)e^{-\frac{x^2+y^2}{2}}\right] \\
&= y(1-x^2)\,e^{-\frac{x^2+y^2}{2}}\cdot -y + (1-x^2)e^{-\frac{x^2+y^2}{2}} \\
&= (1-x^2)(1-y^2)\,e^{-\frac{x^2+y^2}{2}}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial^2 f}{(\partial y)^2} &= \frac{\partial}{\partial y}\left[x(1-y^2)e^{-\frac{x^2+y^2}{2}}\right] \\
&= x\left((1-y^2)\,e^{-\frac{x^2+y^2}{2}}\cdot -y + (-2y)\cdot e^{-\frac{x^2+y^2}{2}}\right) \\
&= -xy(3-y^2)\,e^{-\frac{x^2+y^2}{2}}
\end{aligned}$$

$$\begin{aligned}
D(x,y) &= f_{xx}(x,y)\,f_{yy}(x,y) - f_{xy}(x,y)^2 \\
&= -xy(3-x^2)e^{-\frac{x^2+y^2}{2}}\cdot -xy(3-y^2)e^{-\frac{x^2+y^2}{2}} - \left((1-x^2)(1-y^2)e^{-\frac{x^2+y^2}{2}}\right)^2 \\
&= x^2y^2\,e^{-x^2-y^2}(3-x^2)(3-y^2) - (1-x^2)^2(1-y^2)^2\,e^{-x^2-y^2} \\
&= e^{-x^2-y^2}\left(x^2(3-x^2)y^2(3-y^2) - (1-x^2)^2(1-y^2)^2\right)
\end{aligned}$$

Step 4 (Analyze Crit. Pts):

↑ Use Second Derivative Test

NB: $D(x,y) = D(|x|,|y|)$

$(0,0)$: $D(0,0) = e^{-0}(0-1) = -1 < 0$

Hence $(0,0)$ is a saddle point.

$(1,1)$: $D(1,1) = e^{-2}(1\cdot 2\cdot 1\cdot 2 - 0) = 4e^{-2} > 0$ and $f_{xx}(1,1) = -1\cdot 1\cdot 2\cdot e^{-1} < 0$

Hence $(1,1)$ is a local maximum point

$(1,-1)$: $D(1,-1) > 0$ and $f_{xx}(1,-1) = -1\cdot(-1)\cdot 2e^{-1} = 2e^{-1} > 0$

Hence $(1,-1)$ is a local minimum point.

$(-1,-1)$: $D(-1,-1) > 0$ and $f_{xx}(-1,-1) = -(-1)(-1)\cdot 2e^{-1} = -2e^{-1} < 0$

Hence $(-1,-1)$ is a local maximum point.

$(-1,1)$: $D(-1,1) > 0$ and $f_{xx}(-1,-1) = -(-1)\cdot 1\cdot 2e^{-1} = 2e^{-1} > 0$

Hence $(-1,1)$ is a local minimum point.

//

Absolute Extrema

Prop: Every continuous function $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ with compact domain attains its absolute extreme values.

NB:
- Compact means bounded and closed.
- Compare with the "Closed interval method"...

Ex: Compute the absolute extrema of the function

$f(x,y) = 2x^3 + y^4$ on $D = \{(x,y) : x^2+y^2 \le 1\}$.

↳ is compact



Sol: ① Critical Points: $\nabla f = \langle 6x^2, 4y^3 \rangle = 0 \iff (x,y) = (0,0)$

$\to f(0,0) = 2 \cdot 0^3 + 0^4 = 0$

② Analyze Boundary: parametrize boundary by $\vec{r}(t) = \langle \cos(t), \sin(t) \rangle$ w/ $0 \le t \le \pi$

Consider $g(t) := f(\vec{r}(t)) = 2\cos^3(t) + \sin^4(t)$ on $[0, \pi]$.

$g'(t) = 6\cos^2(t) \cdot -\sin(t) + 4\sin^3(t) \cdot \cos(t)$

$= 2\sin(t)\cos(t)\left(-3\cos(t) + 2\sin^2(t)\right)$

$= \sin(2t)\left(2(1-\cos^2(t)) - 3\cos(t)\right)$

$= -\sin(2t)\left(2\cos^2(t) + 3\cos(t) - 2\right)$

$= -\sin(2t)\left(2\cos(t) - 1\right)\left(\cos(t) + 2\right)$

$g'(t) = 0 \iff \sin(2t) = 0$ OR $\cos(t) = \frac{1}{2}$ ~~OR $\cos(t) = -2$~~

$\iff 2t = 0,\ 2t = \pi,\ 2t = 2\pi,\ 2t = 3\pi,\ 2t = 4\pi,$

$t = \frac{\pi}{3},\ t = \frac{5\pi}{3}$

$\iff t = 0,\ t = \frac{\pi}{3},\ t = \frac{\pi}{2},\ t = \pi,\ t = \frac{3\pi}{2},\ t = \frac{5\pi}{3},\ t = 2\pi$

Critical Points on boundary:

$(1,0)$ [$t=0$], $\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$ [$t=\frac{\pi}{3}$], $(0,1)$ [$t=\frac{\pi}{2}$], $(-1,0)$ [$t=\pi$], $(0,-1)$ [$t=\frac{3\pi}{2}$], $\left(\frac{1}{2}, -\frac{\sqrt{3}}{2}\right)$ [$t=\frac{5\pi}{3}$], $(1,0)$

$f(x,y) = 2x^3 + y^4$

$f\left(\frac{1}{2}, \pm\frac{\sqrt{3}}{2}\right) = \frac{2}{8} + \frac{9}{16} = \frac{13}{16}$

Point	$(0,0)$	$(1,0)$	$\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$	$(0,1)$	$(-1,0)$	$(0,-1)$	$\left(\frac{1}{2}, -\frac{\sqrt{3}}{2}\right)$	$(1,0)$
Value	0	2 max	13/16	1	-2 min	1	13/16	2 max

$\therefore$ f attains absolute maxima at $(1,0)$ and $(0,1)$, and absolute minimum at $(-1,0)$ ☑

Ex: Compute the absolute extrema of $f(x,y) = x + y - xy$ on the triangle with vertices $(0,0)$, $(0,2)$, $(4,0)$.

Sol: Left to you for practice.

Hint 1: Draw a picture of the (filled-in) triangle.

Hint 2: Where is the boundary in your picture?

Hint 3: How do you parametrize a line segment?

Hint 4: Work with one boundary segment at a time.

Ex: Find the points on the cone $z^2 = x^2 + y^2$ which are closest to the point $(4,2,0)$.

Sol: